# LCR

October 4, 2018

## 1 Idea

Assume that the nodes in the graph G have been partitioned into a set of $p$ disjoint clusters $C_1, \cdots, C_p$.

A node $v \in G$ is defined to be a *boundary node* if $v$ is adjacent to another node that belongs to a different cluster from $v$'s cluster.

An edge in $G$ is defined to be a *boundary edge* if its two nodes are boundary nodes.

Let $B_i$ denote the set of boundary nodes in cluster $C_i$.

Given two distinct clusters $C_s$ and $C_t$ in $G$, we say that $C_t$ is reachable from $C_s$ via a $L$-cluster-path $(v_1, v_2, \cdots, v_k)$ if (1) for each $i \in [1, k]$, $v_i \in B_i$; (2) $C_s = C_1$, $C_t = C_k$; and (3) for each $i \in [1, k)$, either (a) $(v_i, v_{i+1})$ is a boundary edge whose label is in $L$ or (b) both $v_i$ and $v_{i+1}$ are boundary nodes in the same cluster (i.e., $C_i = C_{i+1}$), and there is a L-path from $v_i$ to $v_{i+1}$ within $C_i$.

**Lemma 1.1.** Given two distinct nodes $s$ and $t$ in $G$, where $s \in C_s$ and $t \in C_t$, $t$ is reachable from $s$ via a L-path if (1) $C_t$ is reachable from $C_s$ via a $L$-cluster-path $(v_1, v_2, \cdots, v_k)$; (2) $v_1 \in C_s$ is reachable from $s$ via a L-path in $C_s$; and (3) $t$ is reachable from $v_k \in C_t$ via a L-path in $C_t$. $\square$

## 2 Approach

Our approach is based on the three conditions in Lemma 1.1.

To support condition (1), we build a *reachable cluster boundary pairs*, denoted by $RCBP$, where for clusters $C_i$, $C_j$ and $L \subseteq \mathcal{L}$,

$$RCBP(C_i, C_j, L) = \{v_s, v_t : v_s \in B_i, v_t \in B_j, v_s \overset{L}{\rightsquigarrow} v_t\}$$

To facilitate computation, we also build a *reverse reachable cluster index*, denoted by $RRCI$, where for each cluster $C_i$, and $L \subseteq \mathcal{L}$,

$$RRCI(C_i, L) = \{C_j : i \neq j, |RCBP(C_i, C_j, L)| > 0\}$$

To support condition (2), we build a *reachable boundary index* for each cluster $C_i$, denoted by $RBI_i$, where for each node $v \in C_i$ and $L \subseteq \mathcal{L}$,

$$RBI_i(v, L) = \{w \in B_i : v \overset{L}{\rightsquigarrow} w \text{ within } C_i\}$$

To support condition (3), we build a *reverse reachable boundary index* for each cluster $C_i$, denoted by $RRBI_i$, where for each node $v \in C_i$ and $L \subseteq \mathcal{L}$,

$$RRBI_i(v, L) = \{w \in B_i : w \overset{L}{\rightsquigarrow} v \text{ within } C_i\}$$

A LCR query $Q = (s, t, L)$, where $s \in C_s$ and $t \in C_t$, is processed as follows.

1. if $C_s = C_t$ then
   (a) if $(RBI_s(s, L) \cap RRBI_s(t, L) \neq \emptyset)$ then return true
   (b) otherwise, return the outcome determined by a BFS within $C_s$
2. $X = RRCI(C_t, L)$. $X \subseteq \{C_1, \cdots, C_p\}$ denote the clusters that could reach $C_t$ via some L-path in $G$.
3. if $C_s \notin X$, then return false.
4. let $BS = RBI_s(s, L)$. $BS \subseteq B_s$ denote the set of boundary nodes in $C_s$ reachable from s via some L-path within $C_s$.
5. if $BS = \emptyset$, then return false.
6. let $BT = RRBI_t(t, L)$. $BT \subseteq B_t$ denote the set of boundary nodes in $C_t$ that could reach t via some L-path within $C_t$.
7. if $BT = \emptyset$, then return false.
8. $XV = RCBP(C_s, C_t, L)$. $XV = \{v_i, v_j\}$, $v_i \subseteq B_s$, $v_j \subseteq B_t$, denote the pairs of boundary nodes that connects cluster $C_i$ and $C_j$.
9. if there exists a pair of $(v_i, v_j)$, $v_i \in BS$, $v_j \in BT$, $(v_i, v_j) \in XV$, then return true.
10. return false

# 3 Space & Time Complexity

Let the nodes be partitioned into $p$ clusters, $n_i$ denote the number of nodes in cluster $C_i$, and $b_i$ denote the number of boundary nodes in $C_i$.

The number of entries in RCBP is $O(2^{\mathcal{L}} \sum_{1 \leq i,j \leq p} b_i b_j)$. The number of entries in RRCI is $O(p^2 2^{\mathcal{L}})$. The number of entries in RBI/RRBI is $O(2^{\mathcal{L}} \sum_{i=1}^{p} n_i b_i)$.

In the worst case, all possible pairs of boundary nodes $(v_s, v_t), v_s \in B_s, v_t \in B_t$ are checked against $RCBP(C_s, C_t, L)$.